

Comparing and Contrasting Depression Screening Instruments for Use Among Adolescents in Primary Care

Mary Clinton¹, SN-CCC
Julie Kaszuba², MSN, RN

¹ *Research Scholar-LVPG Clinical Services, Student Nurse at Cedar Crest College*

² *LVPG Clinical Procedure & Products Nurse Specialist, Research Scholar Mentor*

Abstract

This discussion encompasses 18 research articles focusing on instruments used for depression screening of adolescent patients. Early detection of depression is a crucial step for stopping the progressive course of a depressive disorder. Effective screening mandates a reliable, valid, and brief tool is used. Lehigh Valley Physician Group (LVPG) is seeking to identify and implement a depression screening instrument aligned with the Lehigh valley health network (LVHN) Triple Aim. Better care, better cost, and better health for adolescent patients in primary care can be measured post implementation of a standard depression screening instrument. Four depression screening instruments and one set of guidelines were reviewed. The Patient Health Questionnaire-9 (PHQ-9), the Patient Health Questionnaire for Adolescents (PHQ-A), the Beck Depression Inventory-II (BDI-II), the Beck Depression Inventory for Primary Care (BDI-PC), and the Guidelines for Adolescent Preventive Services (GAPS) are evaluated. Evidence illustrates the instruments have comparable characteristics, including language translations, statistical analysis, and gender bias, but differ on cost, purpose of design, and time to administer. Collaboration with subject matter experts prior to use of a standard depression screening instrument in primary care is recommended.

Keywords: *adolescence, adolescent depression, depression screening, Patient Health Questionnaire, PHQ, Patient Health Questionnaire-9 Item, PHQ-9, Patient Health Questionnaire for Adolescents, PHQ-A, Beck Depression Inventory, BDI, Beck Depression Inventory for Primary Care, BDI-PC, Beck Depression inventory-II, BDI-II, Guidelines for Adolescent Preventive Services, GAPS, primary care, and standardization*

Background

Adolescence is a time of emotional, physical, and mental growth. In the midst of this transformational period, an estimated 30% of children age 12 to 17 are more prone to depressive symptoms (Dolle et al., 2012).

Depression prevention is vital to improving a patient's quality of life. "Depressive illness is projected to have significant public health and economic costs: major depression is expected to be the second leading cause of death and disability and to impose the greatest burden of ill health worldwide by 2020" (Huang, Chung, Kroenke, Delucchi, & Spitzer, 2006, p. 547). The Centers for Medicaid and Medicare Services (CMS) and The U.S. Preventive Services Task Force (USPSTF) encourage annual depression screenings. Notably, "The U.S. Preventive Services Task Force (USPSTF) recommends screening adolescents (12 to 18 years of age) for major depressive disorder (MDD) when systems are in place to ensure accurate diagnosis, psychotherapy (i.e., cognitive behavioral or interpersonal) and follow up" (U.S. Preventive Services Task Force, 2010, p. 178).

instrument that accurately detects adolescent depression in primary care settings. Improved patient outcomes can be achieved by applying measures reflecting better cost, better care, and better health to patients served. Selecting a depression screening instrument that is reliable, valid, and "should ideally have both a high sensitivity and a high specificity in order to reduce the number of false-negatives and false-positives" is necessary (Wittkamp, Naeije, Schene, Huyser, & van Weert, 2007, p. 388). Effective instruments, "must be valid, reliable, brief, and easy to use" (Gilbody, Richards, Brealey, & Hewitt, 2007, p.1596).

Four depression screening instruments are evaluated in this paper. These include the Patient Health Questionnaire-9 Item (PHQ-9), the Patient Health Questionnaire for Adolescents (PHQ-A), the Beck Depression Inventory-II (BDI-II) and the Beck Depression Inventory for Primary Care (BDI-PC). In addition, the usefulness of the Guidelines for Adolescent Preventive Services (GAPS) in screening for adolescent depression is critiqued.

Lehigh Valley Physician Group (LVPG) is investigating an age appropriate depression screening

Methodology

From June 8, 2015-July 13, 2015, literature addressing instruments used for adolescent depression screening instruments were retrieved for analysis. Databases searched include: CINAHL, HAPI, Medline, PubMed, EBSCO, Pediatrics, and Science Direct. Key search terms included: *adolescence, adolescent depression, depression, depression screening, depression measurement, mood module, Patient Health Questionnaire, PHQ, Patient Health Questionnaire-9 Item, PHQ-9, Patient Health Questionnaire for Adolescents, PHQ-A, Beck Depression Inventory, BDI, Beck Depression Inventory for Primary Care, BDI-PC, Beck Depression inventory-II, BDI-II, Guidelines for Adolescent Preventive Services, and GAPS.*

Initially, search settings were not placed for patient race or ethnicity, culture, type of care setting, nor age of patients aged 12 to 17 years. Limiting age and type of care setting searched was necessary to acquire additional evidence specifically aimed at screening for adolescent depression. Two Lehigh Valley Health Network (LVHN) medical librarians were consulted as experts for refining the search. Furthermore, this author collaborated with the LVPG Clinical Quality Educator to review data reflecting compliance with annual depression screening in primary care. Subsequent discussions emphasized the need for standardized annual depression screening in LVPG primary care practices. An evidence table was constructed. Rating the level of evidence assisted in identifying the most valid and reliable data. This author and mentor met weekly to review research and examine findings.

Review of Literature

The PHQ-9 is a self-administered mood module. The contents of this depression screening instrument are extracted from the Patient Health Questionnaire (PHQ), which is a self-administered version of the Primary Care Evaluation of Mental Disorders (PRIME-MD) that detects depressive and mental disorders in primary care (Wittkamp et al., 2007). The instrument “consists of nine items taken directly from the depression criteria” listed in the Diagnostic and Statistical Manual Fourth Edition (DSM-IV) (Kung et al., 2013, p.341). It is used for the adult population in a variety of medical settings and is the current standard adult depression screening instrument in LVPG primary care practices. Patients rate symptoms experienced over a period of two weeks prior to administration of the screening. In particular, one question assesses difficulty

performing tasks (Kroenke, Spitzer, & Williams, 2001). Scoring the instrument is completed by physicians using either a diagnostic algorithm or a recommended cut-off score (Kroenke et al., 2001). Although the PHQ-9 can be scored using either method, the algorithm method has reported a low sensitivity and is not recommended. “A cut-off score of 10 or above on the summed-item score has been recommended as a method for screening for major depressive disorder” (Manea, Gilbody, & McMillan, 2015, p. 68). Calculating scores using a method reporting a higher sensitivity meets the criteria for a depression screening instrument described by Wittkamp et al. (2007).

The PHQ-A is an instrument specifically constructed to screen the target age range, patients age 12 to 17. This instrument is a “67 item questionnaire that can be entirely self-administered by the patient in 5 minutes or less” and was “developed for the assessment of mental disorders among adolescent primary care patients” (Johnson, Harris, Spitzer, & Williams, 2002, p. 197). Answers are scored using a diagnostic algorithm. The PHQ-9 Modified for Adolescents (PHQ-A), has been extracted from the full version of the PHQ-A to screen for adolescent depression according to DSM-IV criteria. Symptoms experienced by the patient in the two weeks prior to screening are measured. Like the PHQ-9, the PHQ-A includes a question that measures functional impairment. It also inquires about suicidal ideation and suicide attempts (Johnson et al., 2002).

Evidence demonstrates it is common for brief depression screening instruments to evolve from original, lengthier, screeners. Following this trend, the BDI-II was developed in the late 1990’s from the Beck Depression Inventory (BDI) to detect depressive symptoms as listed in the DSM-IV (Kung et al., 2013). Considering its revision, the BDI is not discussed. The BDI-II consists of 21 questions and has proven to be valid for detecting depression in patients 13 and older (Dolle et al., 2012). Seven questions from the BDI-II were taken to create the BDI-PC, a self-administered screening instrument for primary care that places greater importance on assessing a patient’s affect and mentality. The instrument uses a scale to gauge severity of depression according to DSM-IV criteria (Winter, Steer, Jones-Hicks, & Beck, 1999). Dolle et al. report the BDI-II has patients rate symptoms experienced in the last 14 days, including the day of the screening (2012). The BDI-PC states the same timeframe for rating symptoms experienced (Steer, Cavalieri, Leonard, & Beck, 1999). Reviewing the American Academy of Family Physicians’ (AAFP)

recommendations for depression screening instruments across the lifespan, Sharp and Lipsky noted neither the BDI-II nor the BDI-PC was designated for use in the target population, although the BDI was recommended (2002).

While the above questionnaires are defined as depression screening instruments, a set of guidelines, the GAPS, has been listed as resource for primary care. GAPS contains a series of questionnaires used to assess adolescent risk, including depressed mood (Gadomski, Scribani, Krupa, & Jenkins, 2014). Evidence illustrates the GAPS' were designed to assist healthcare professionals in providing immunizations, conducting annual screenings, detecting behavioral problems, and promoting healthy living among adolescents (Levenberg, 1998). Thus, this set of recommendations has been eliminated from this discussion based on the absence of supporting data.

Cost According to Kung et al., there is no cost to administer the PHQ-9 (2013). Furukawa noted the BDI-II must be purchased due to copyright (2010). A single screen costs about two dollars (Kung et al., 2013). The BDI-PC must also be purchased, although no cost was listed (Sharp & Lipsky, 2002). It is not determined whether the PHQ-A is free for use or has to be purchased.

Time Johnson et al. report the PHQ-A takes approximately five minutes to complete (2002). Completing and scoring the PHQ-9 takes less than two minutes (Furukawa, 2010). Different times have been listed for the completion of the BDI-II. Although it consists of 21 questions, the BDI-II can be completed by the patient in a short amount of time. According to Kung et al., the instrument can be answered in five minutes (2013). In comparison, Furukawa states the 21 questions from the BDI-II can be completed in a minimum of five minutes but may take up to ten minutes (2010). Reported by the AAFP, the BDI-PC is completed in "fewer than 5 minutes" (Sharp & Lipsky, 2002).

Language According to the AAFP, the BDI-II, and BDI-PC are available in Spanish (Sharp & Lipsky, 2002). The PHQ-9 has been translated into several languages, including German, Portuguese, Thai, Dutch, Malay, and Konkani (Manea et al., 2012). Johnson et al. did not report if any translated versions of the PHQ-A exist (2002).

Reading Level No reading levels for any of the instruments was reported.

Race/Ethnicity Among a primary care population, no relationship was identified between BDI-II scores and race or ethnicity among "Caucasian", "African American", "Hispanic", "Asian American/Pacific Islander", "Other," and "Unreported" primary care patients (Arnau, Meagher, Norris, & Bramson, 2001, p. 113-114). Also proven useful for screening among a diverse patient population, the BDI-PC scores were not influenced by a patient's ethnicity (Winter et al., 1999). After administering the PHQ-9 to a group of primary care patients, it was concluded "in African American, Chinese American, Latino, and non-Hispanic white patient groups the PHQ-9 measures a common concept of depression and can be effective for the detection and monitoring of depression in these diverse populations" (Huang et al., 2006, p. 547).

Gender Bias No gender bias was reported in the literature gathered for evaluating the PHQ-A. More women than men were diagnosed with depression using the PHQ-9, BDI-II, and BDI-PC.

Sensitivity and Specificity Statistics were compared among studies using the cutoff value of 10 and higher for the PHQ-9. In one meta-analysis, the PHQ-9 revealed a sensitivity of 0.77 and a specificity of 0.94 (Wittkamp et al., 2007). It was not identified whether this data represented an adolescent, adult, or mixed population. Measuring sensitivity and specificity, a second meta-analysis found the PHQ-9 to have a sensitivity of 0.85 and specificity of 0.89 at a cutoff of 11. In this case, the screening instrument was administered to an adult population (Manea et al., 2012).

Screening for depression in an adolescent population, a sensitivity of 0.895 and specificity 0.775 was recorded at a cutoff score of 11 (Richardson et al., 2010).

In comparison to the PHQ-9, the PHQ-A identified a sensitivity of 0.73 and a specificity of 0.94 in adolescents with major depressive disorder (Johnson et al., 2002).

The BDI-II has measured a sensitivity 1.0 and specificity 0.70 at a cutoff of 10 in a primary care population (Arnau et al., 2001). The BDI-PC was found to have sensitivity 0.91 and specificity 0.91 in adolescent medical outpatients (Winter et al., 1999). A sensitivity of 0.97 and specificity of 0.99 was reported for a cutoff of 11 in an adult primary care population (Steer et al., 1999).

See Table 1 for details: Depression Screening Instrument Statistics

Area Under the Curve (AUC) Assessing validity of the PHQ-9 for diagnosing major depression in adults revealed the PHQ-9 to have an area under the curve of 0.95 (Kroenke et al., 2001). Conducting a study among an adolescent population, Richardson et al. discovered the PHQ-9 to have an AUC of 0.88 at a cutoff of 11 (2010). The BDI-II was found to have an accuracy of 0.96 in an adult primary care population (Arnau et al., 2001). The instrument also has an AUC of 0.93 when used among German adolescent mental health patients (Dolle et al., 2012). Winter et al. discovered the BDI-PC has an AUC of 0.98 when screening is conducted among adolescent medical outpatients (1999). At a cutoff of 4, Steer et al. reported an AUC of 0.99 for the same instrument when it was used in an adult primary care population (1999).

See Table 1 for details: Depression Screening Instrument Statistics

Discussion

The PHQ-A, PHQ-9, BDI, and BDI-PC have proven to be valid and reliable instruments available to screen for adolescent depression. Having comparable completion times and statistics gives health networks various instruments to choose from. Translated versions of the BDI-II, BDI-PC, and the PHQ-9 do exist. Each of these screening instruments has been used among patients of different race and culture. Each instrument reports higher rates of depression in females.

A few differences exist among these instruments. The full version of the PHQ-A takes more time to complete. Nine of its questions assessing depression have been placed into a questionnaire, the PHQ-9 Modified for Adolescents (PHQ-A), to reduce completion time. The BDI-II takes several more minutes to complete than the other instruments. Free to use, the PHQ-A and PHQ-9 may be preferable since the screening is performed annually. The BDI-II and BDI-PC must be purchased. The benefit of using the PHQ-A is its development for an adolescent population and inclusion of a question about suicidal ideation and suicide attempts. Although it was not designed specifically for adolescents, the PHQ-9 is the current standard depression screening instrument for adults in LVPG primary care. Medical professionals are already familiar with this depression screening instrument and may require less time to develop competency when applying the PHQ-9 to the adolescent population. However, one might argue that although medical professionals may be more familiar with the PHQ-9, the BDI-PC has

fewer questions. This instrument would take less time to complete and score.

Limitations

While database searching proved to be thorough, a minimal amount of literature was found using the PHQ-A, PHQ-9, BDI-II, and BDI-PC depression screening instruments in adolescents. Only one article, by the instrument's creator, was found to report the use of the PHQ-A in adolescents. A significant amount of articles were found for use of the BDI rather than the BDI-II or BDI-PC, despite these being updated versions of the BDI. Although the BDI-II has been the updated version of the BDI since the late 1990's, there leaves to be some speculation why the original BDI is recommended for use in adolescents by the AAFP (Sharp & Lipsky, 2002).

Recommendations

Consulting subject matter experts, including adolescent health, behavioral health, and pediatric healthcare professionals is advised before implementing a standard adolescent depression screening instrument in LVPG primary care. Collaborating with these healthcare professionals can provide additional insight into providing quality care for the target age range, patients age 12 to 17. Expert advice, time restraints, and languages spoken by the population served, among other factors, may influence the selection of the instrument chosen for LVPG primary care.

References

- Arnau, R. C., Meagher, M. W., Norris, M. P., & Bramson, R. (2001). Psychometric evaluation of the Beck Depression Inventory-II with primary care medical patients. *Health Psychology: Official Journal of the Division of Health Psychology, American Psychological Association*. 20(2), 112-119.
- Dolle, K., Schulte-Korne, G., O'Leary, A. M., von Hofacker, N., Izat, Y., & Allgaier, A. K. (2012). The Beck Depression Inventory-II in adolescent mental health patients: Cut-off scores for detecting depression and rating severity. *Psychiatry Research*. 200(2-3), 843-848.
- Furukawa, T. A. (2010). Assessment of mood: Guides for clinicians. *Journal of Psychosomatic Research*. 68(6), 581-589.

Gadomski, A. M., Scribani, M. B., Krupa, N., & Jenkins, P. L. (2014). Do the Guidelines for Adolescent Preventive Services (GAPS) facilitate mental health diagnosis? *Journal of Primary Care & Community Health*. 5(2), 85-89.

Gilbody, S., Richards, D., Brealey, S., & Hewitt, C. (2007). Screening for Depression in Medical Settings with the Patient Health Questionnaire (PHQ): a Diagnostic Meta-Analysis. *Journal of General Internal Medicine*. 22(11), 1596-1602.

Huang, F. Y., Chung, H., Kroenke, K., Delucchi, K.L., & Spitzer, R. L. (2006). Using the Patient Health Questionnaire-9 to measure depression among racially and ethnically diverse primary care patients. *Journal of General Internal Medicine*. 21(6), 547-552.

Johnson, J. G., Harris, E. S., Spitzer, R. L., & Williams, J. B. W. (2002). The Patient Health Questionnaire for Adolescents - Validation of an instrument for the assessment of mental disorders among adolescent primary care patients. *Journal of Adolescent Health : Official Publication of the Society for Adolescent Medicine*. 30(3), 196-204.

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9 Validity of a Brief Depression Severity Measure. *Journal of General Internal Medicine*. 16(9), 606-613.

Kung, S., Alarcon, R. D., Williams, M. D., Poppe, K. A., Jo Moore, M., & Frye, M. A. (2013). Comparing the Beck Depression Inventory-II (BDI-II) and Patient Health Questionnaire (PHQ-9) depression measures in an integrated mood disorders practice. *Journal of Affective Disorders*. 145(3), 341-343.

Levenberg, P. B. (1998). GAPS: an opportunity for nurse practitioners to promote the health of adolescents through clinical preventive services. *Journal of Pediatric Health Care : Official Publication of National Association of Pediatric Nurse Associates & Practitioners*. 12(1), 2-9.

Manea, L., Gilbody, S., & McMillan, D. (2012). Optimal cut-off score for diagnosing depression with the Patient Health Questionnaire (PHQ-9): a meta-analysis. *CMAJ : Canadian Medical Association Journal = Journal De L'Association Medicale Canadienne*. 184(3), E191-E196.

Manea, L., Gilbody, S., & McMillan, D. (2015). A diagnostic meta-analysis of the Patient Health Questionnaire-9 (PHQ-9) algorithm scoring method as a screen for depression. *General Hospital Psychiatry*. 37(1), 67-75.

Richardson, L. P., McCauley, E., Grossman, D. C., McCarty, C. A., Richards, J., Russo, J. E., Rockhill, C., & Katon, W. (2010). Evaluation of the Patient Health Questionnaire-9 Item for Detecting Major Depression Among Adolescents. *Pediatrics*. 126(6), 1117-1123.

Sharp, L. K., & Lipsky, M. S. (2002). Screening for depression across the lifespan: a review of measures for use in primary care settings. *American Family Physician*. 66(6), 1001-8.

Steer, R. A., Cavalieri, T. A., Leonard, D. M., & Beck, A. T. (1999). Use of the Beck Depression Inventory for Primary Care to screen for major depression disorders. *General Hospital Psychiatry*. 21(2), 106-11.

U.S. Preventive Services Task Force Screening and treatment for major depressive disorder in children and adolescents: Recommendation statement. (2010). *American Family Physician*. 82(2), 178-179.

Winter, L. B., Steer, R. A., Jones-Hicks, L., & Beck, A. T. (June 01, 1999). Screening for major depression disorders in adolescent medical outpatients with the Beck Depression Inventory for Primary Care. *Journal of Adolescent Health*, 24(6), 389-394.

Wittkamp, K. A., Naeije, L., Schene, A. H., Huyser, J., & van Weert, H. C. (2007). Diagnostic accuracy of the mood module of the Patient Health Questionnaire: a systematic review. *General Hospital Psychiatry*. 29(5), 388-395.

Depression Screening Instrument	Sensitivity	Specificity	AUC
PHQ-A	0.73	0.94	-
PHQ-9	0.77 0.85 (Cutoff 11) 0.895 (Cutoff 11)	0.94 0.89 (Cutoff 11) 0.775 (Cutoff 11)	0.95 0.88 (Cutoff 11)
BDI-II	1.0 (Cutoff 10)	0.7 (Cutoff 10)	0.96 0.93
BDI-PC	0.91 0.97	0.91 0.99	0.98 0.99

Table 1. Depression Screening Instrument Statistics

References

Arnau, R. C., Meagher, M. W., Norris, M. P., & Bramson, R. (2001). Psychometric evaluation of the Beck Depression Inventory-II with primary care medical patients. *Health Psychology : Official Journal of the Division of Health Psychology, American Psychological Association*. 20(2), 112-119.

Dolle, K., Schulte-Korne, G., O'Leary, A. M., von Hofacker, N., Izat, Y., & Allgaier, A. K. (2012). The Beck Depression Inventory-II in adolescent mental health patients: Cut-off scores for detecting depression and rating severity. *Psychiatry Research*. 200(2-3), 843-848.

Johnson, J. G., Harris, E. S., Spitzer, R. L., & Williams, J. B. W. (2002). The Patient Health Questionnaire for Adolescents - Validation of an instrument for the assessment of mental disorders among adolescent primary care patients. *Journal of Adolescent Health : Official Publication of the Society for Adolescent Medicine*. 30(3), 196-204.

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9 Validity of a Brief Depression Severity Measure. *Journal of General Internal Medicine*. 16(9), 606-613.

Richardson, L. P., McCauley, E., Grossman, D. C., McCarty, C. A., Richards, J., Russo, J. E., Rockhill, C., & Katon, W. (2010). Evaluation of the Patient Health Questionnaire-9 Item for Detecting Major Depression Among Adolescents. *Pediatrics*. 126(6), 1117-1123.
Manea, L., Gilbody, S., & McMillan, D. (2012). Optimal cut-off score for diagnosing depression with the Patient Health Questionnaire (PHQ-9): a meta-analysis. *CMAJ : Canadian Medical Association Journal = Journal De L'Association Medicale Canadienne*. 184(3), E191-E196.

Steer, R. A., Cavalieri, T. A., Leonard, D. M., & Beck, A. T. (1999). Use of the Beck Depression Inventory for Primary Care to screen for major depression disorders. *General Hospital Psychiatry*. 21(2), 106-11.

Winter, L. B., Steer, R. A., Jones-Hicks, L., & Beck, A. T. (June 01, 1999). Screening for major depression disorders in adolescent medical outpatients with the Beck Depression Inventory for Primary Care. *Journal of Adolescent Health*, 24(6), 389-394.

Wittkamp, K. A., Naeije, L., Schene, A. H., Huyser, J., & van Weert, H. C. (2007). Diagnostic accuracy of the mood module of the Patient Health Questionnaire: a systematic review. *General Hospital Psychiatry*. 29(5), 388-395.